A New Look at the Forward Premium "Puzzle"

HAITHAM A. AL-ZOUBI*

We decompose the spot and forward rates into (permanent) nonlinear trend components and (transitory) stationary components. We examine the unbiasedness of the permanent (transitory) component of the forward rate in predicting the permanent (transitory) component of its corresponding future spot rate. The transitory component of the future spot rate under reacts to the transitory component of the forward rate. However, the permanent component of the forward rate is an unbiased predictor of the permanent component of the future spot rate. A robust nonlinear cotrending relation is also found between the forward and future spot rates and the hypothesis of the forward-rate unbiasedness is sustained in the long run. These results suggest that the forward rate better explains the long-term behavior of the future spot rate. Simulation analysis shows that if the transitory component of the forward rate fully predicts the transitory component of the future spot rate, the forward premium puzzle disappears. © 2010 Wiley Periodicals, Inc. Jrl Fut Mark 31:599–628, 2011

INTRODUCTION

The hypothesis that forward exchange rates are unbiased predictors of future spot rates is empirically far from conclusive. Theoretically, forward rates conditional bias is attributed to two broad categories: (1) the presence of forecast errors and (2) the existence of time-varying risk premium.

The forecast errors category has several explanations for the forward biasedness. First, the peso problem of Krasker (1980), where there are sustained

We are very grateful to Herman Bierens and Gerald Whitney for their comments and suggestions.

*Correspondence author, Department of Economics and Finance, United Arab Emirates University, 17555, Al-Ain, Abu Dhabi, United Arab Emirates. Tel: +971-50-1383553, Fax: +971 7624384.

Received May 2010; Accepted August 2010

■ *Haitham A. Al-Zoubi is an Associate Professor of Finance, United Arab Emirates University.*

The Journal of Futures Markets, Vol. 31, No. 7, 599–628 (2011)

Published online October 8, 2010 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/fut.20486

excess forward premia for a long period of time resulting from investors' belief of low likelihood of large depreciation. Second, the learning problem of Lewis (1995), by which investors adapt their expectations slowly to shifts in macroeconomic measurable fundamentals. Third, the irrational investors model suggested by Frankel and Froot (1990), where the behavior of noise traders "chartists" interacts with rational agents to produce potentially a long movement in exchange rates and a failure of uncovered interest parity.[1] Cavaglia, Verschoor, & Wolf (1993) conclude that the failure of forward rates unbiasedness is attributed to both irrationality of expectations and the time-varying risk premia.

Others attribute the findings of forward bias to time-varying risk premium. (e.g. Fama, 1984; Hansen & Hodrick, 1983; Hsieh, 1989). However, Crowder (1994) finds that risk premium (proxied by forward premium) behaves like a non-stationary stochastic process. Since the cointegrating vector, by definition, is stationary, forward premium fail to identify risk premium.[2] Engle (1996) concludes, "Models of risk premium have been unsuccessful at explaining the magnitude of this failure of unbiasedness" (p. 124).

In this article, however, we will take a different route. We focus primarily on the ability of the forward rates to track spot rates movements over short and long horizons. Prior studies find weak evidence, at best, that macroeconomic variables—such as interest rates, prices, and GDP—can forecast short-term movements of spot rates. In their seminal paper, Meese and Rogoff (1983) conclude that while macroeconomic measurable fundamentals can explain exchange rate changes over medium and long horizons, they are not useful for tracking spot rate changes over short horizons. Evans and Lyons (2005) compare the true, ex-ante forecasting performance of a micro-based model against both Meese and Rogoff macro model and random walk. They conclude that the micro-based model suggested by Engle and West (2004, 2005) constantly outperforms random walk and the macro model.[3]

We argue that the rejection of the forward-rate unbiasedness hypothesis to the failure of the transitory component of the forward rate to fully predicts the transitory component of the future spot rate. We conclude that the forward rate is poor in tracking spot rate movements over short horizons. However, we show that the permanent component of the forward rate, which we model as a nonlinear deterministic trend can fully predict the nonlinear deterministic trend component of the corresponding future spot rate. Evidence of a one-to-one

[1]Engle and Hamilton (1990) show that the value of the dollar tends to move for a long period of time, then we could expect that the forward premia persist for a long time horizon.
[2]Hodrick (1987), Nijman, Palm, and Wolff (1993), and Bams, Walkowiak, and Wolff (2004) demonstrate that low-order autoregressive models can mimic the true pattern of risk premia quite adequately.
[3]The asset approach to exchange rates proposed by Engle and West (2004, 2005) suggests that in rational expectation present-value model, the exchange rate follows near-random walk behavior if there exist some unobserved fundamental that itself follows a random walk.

cotrending relation is found between the forward rate and the future spot rate. Furthermore, we show that the magnitude of the transitory components of the spot and forward rates is responsible for the forward premium puzzle. Monte Carlo simulations show that if the transitory component of the forward rate were an unbiased predictor of the transitory component of the future spot rate, then the puzzle disappears.

The motivation for a relation between nonlinear trend components of the spot and forward rates is threefold. First, although the main line of empirical research models the spot and forward rates as random walk processes,[4] many researchers provide evidence that foreign exchange spot and forward rates are mean reverting. (e.g. Huizinga, 1987; Nikolaou, 2008; Rogoff, 1996; Sollis, 2008; Wu & Chen, 1998). However, the spot and forward rates still behave like cointegrated process in that the series shared a common trend, which is only probable for random walk process. A potential clarification is that the forward and spot rates have frequent nonlinear deterministic time trends.

Second, empirical research documents that exchange rate returns are predictable; they are positively correlated over some few months (bandwagon effect and momentum) and negatively correlated over longer horizons.[5] The predictability of exchange rate returns raises the question of whether market transactions should be considered as stochastic events, deterministic events, or a mixture of both. Certainly, investors will respond to fundamental events, and to the extent that these signals are stochastic, the response of the traders will be stochastic as well. However, investors' reactions are unlikely to be fully automatic. The realized forward premium will reflect both arrivals of new information, which might be stochastic, and the subjective assessment of these stochastic information by traders, which may be considered as deterministic and time-varying. Such behavior is in accordance with all equilibrium asset pricing models, which assume time-varying risk premium. In equilibrium asset pricing framework, the consumption smoothing is the driving force behind the relation. In bad times, when consumption is low (close to habit), investors have high degree of risk aversion and require higher risk premium (see Campbell and Cochrane, 1999; Epstein and Zen, 1989; Verdelhan, 2010, for example).

Verdelhan (2010) develops a habit model that explains the forward anomaly. In his model, the risk aversion is countercyclical and interest rates are

[4]See for example Roll (1979), Frankel (1981), Alder and Lehman (1983), Hakkio (1984, 1986), Corbae, Lim, and Oulairis (1992), and Phillips and McFarland (1997).

[5]Sims (1988) argues that the use of the conventional unit root tests on foreign exchange data are biased toward the acceptance of the random walk hypothesis if the size of the autocorrelation is large. As an alternative, he introduces a Basian test that can discriminate between the random walk and the slowly mean reverting fads behavior suggested by Summers (1986). He strongly rejects the random walk hypothesis in the data-generating process. Using long horizons autoregressions, Huizinga (1987) shows that exchange rate returns are negatively correlated. The mean reversion evidence is also reported in subsequent studies, such as Abuaf and Jorion (1990) and Whitt (1992).

procyclical. In bad times, when risk aversion is high and domestic interest rates are low, investors require positive currency excess returns. Fama and French (1993) and Lettau and Ludvigson (2009) provide an ample evidence that risk premia move inversely with business cycle. Ang, Bekaert, and Wei (2008) and Clarida, Gali, and Gertler (2000) conclude that the interest rates are procyclical. Conrad and Kaul (1998), Berk, Green, & Naik (1999), and Chordia and Shivakumar (2002) provide evidence that time-varying expected returns being a plausible explanation for stock momentum. To the extent that the predictability of stock returns by macroeconomic fundamentals is due mainly to the ability of these fundamentals to capture time-varying risk premium. Such behavior assumes that stock returns have a time-varying unconditional expectation.

Engle and Hamilton (1990) reject the hypothesis that foreign exchange rates follow a random walk in favor of a sequence of stochastic, segmented time trends. They show that past values of the changes in the exchange rates help predict the future in a nonlinear manner. They model the exchange rates as discrete-state first-order Markov process, assuming that the transition probabilities are constant, which eliminates the heterogeneity. However, Hamilton's approach does not solve for the momentum anomaly; it only shifts it to another level: are the transition probabilities time-invariant, or not? If not, then it is reasonable to assume that the forward premium have a time-varying unconditional expectation.

Finally, there are now theoretical explanations of nonlinearities in the exchange rates arising from: the incidence of substantial transaction costs resulting from eliminating arbitrage opportunities in the financial markets (Baldwin, 1990), investors' arbitrage to eliminate capital imbalances across countries (Dumas, 1992; Hollified & Uppal, 1997), the presence of limits to speculation (Lyons, 2001), the presence of heterogeneous traders and strong reversion to fundamentals (Kilian & Taylor, 2003, and sporadic central banks interventions (Mark & Moh, 2004), and the countercyclical behavior of risk premium (Verdelhan, 2010).

There has been some empirical research, which has begun to deal with the nonlinear dynamics of the exchange rates. In particular, Balke and Fomby (1997), Taylor and Peel (2000), Taylor, Peel, & Sarno (2001), and Kilian and Taylor (2003) provide evidence of various nonlinearities in deviation of exchange rates from fundamentals.[6] A growing amount of research is directed

[6]Taylor and Peel (2000) provide evidence that the relationship between the spot exchange rate and monetary fundamentals is nonlinear, given that monetary fundamentals are clearly linked to the relative prices. Obtsfeld and Taylor (1997) show that the deviations of the exchange rates from the low of one price have nonlinear dynamics. Using the exponential smooth transition autoregressive (ESTAR) model of TerÄrsvirta (1994), Taylor et al (2001) provide evidence of nonlinear dynamics in the real exchange rates. Kilian and Taylor (2003) also document the ESTAR dynamics on the long horizon predictability of nominal exchange rates.

at the issue. Some authors report convincing evidence of nonlinear asymmetric mean reversion in real exchange rates (Nikoaou, 2008). Using different approach, Yang, Su, and Kolari (2008) find nonlinear predictability in terms of economic criteria in the martingale behavior of the exchange rates. Additionally, Smallwood (2008) finds both long memory and threshold nonlinearity in the dynamics of real exchange rates. Utilizing time-varying smooth transition autoregressive (TV-STAR) models, Sollis (2008) reports both nonlinearity and structural change. Some suggest that the forward bias documented in the literature may be less indicative of major market inefficiencies. For example, Sarno, Valente, & Leon (2006) provide evidence that deviations from the uncovered interest rate parity display significant nonlinearities, consistent with Lyons (2001) limits to speculation theory. Other researchers report a nonlinear response that is directly related to the current state of the economy. For example, Bansal (1997) and Bansal and Dahlquist (2000) find that the sign of the estimated slope coefficient in the forward premium regression is strongly related to the sign of the interest rate differential. The forward premium anomaly is more probable when the US interest rates are less than foreign rates.

The rest of the paper is organized as follows. Section 2 presents the econometric specification of the nonlinear deterministic trend and the nonlinear cotrending procedure that we consider. Section 3 describes the data and its time series properties. In Section 4, the nonlinear cotrending relation between the forward rate and the future spot rate is examined and test the unbiasedness hypothesis is conducted. Section 5 shows that the stationary components (not the nonlinear trend components) of spot and forward rates are responsible for the routinely rejection of the expectations hypothesis. Section 6 analyzes the consequence of our empirical findings for the forward premium puzzle.

NONLINEAR TREND STATIONARY AND NONLINEAR COTRENDING

To motivate the nonlinear cotrending test of Bierens (2000), consider the kind of nonlinear trend stationary process of the following form:

$$z_t = g(t) + u_t, \tag{1}$$

where $g(t) = \beta_{0,n} + \beta_{1,n}t + f_n(t)$, z_t is a k-variate time series process, u_t is a k-variate zero mean stationary process, and $f_n(t)$ is some deterministic k-variate nonlinear trend function.[7]

[7]We consider two cases for the nonlinear deterministic trend, the first where $f(t)$ is constant except for a single jump and the second where it is piecewise linear in time with connected adjacent pieces.

The hypothesis of nonlinear cotrending is that: assuming the time series are stationary about nonlinear deterministic time trends, nonlinear cotrending is the phenomenon that one or more linear combinations of the time series are stationary about a constant or linear time trend, hence there exists a non-zero vector θ such that $\theta^T f_n(t)$ is orthogonal to zero. Tests of nonlinear cotrending are concerned with testing the null hypothesis $\Omega(r)$ that the space of all vectors θ has dimension r, against the alternative hypothesis $\Omega(0)$ that this dimension is zero.

It is suggested that $f_n(t)$ is the OLS residuals of the regression of $g(t)$ on an intercept and time t and that

$$\sum_{t=1}^{n} f_n(t) = 0, \ \sum_{t=1}^{n} t\, f_n(t) = 0. \tag{2}$$

Bierens (2000) suggests to consider the common eigenvector of the matrices

$$\hat{M}_1 = (1/n) \sum_{t=1}^{n} \hat{F}(t/n)\hat{F}(t/n)^T \tag{3a}$$

where

$$\hat{F}(x) = (1/n) \sum (z_t - \hat{\beta}_0 - \hat{\beta}_1 t) \ \textit{if}\ x \in [n^{-1}, 1], \hat{F}(x) = 0 \text{ if } x \in [0, n^{-1}] \tag{3b}$$

and

$$\hat{M}_2 = (1/n) \sum_{t=1}^{n} \hat{F}'\ (t/n)\ \hat{F}'(t/n)^T \tag{4a}$$

where

$$\begin{aligned} &\hat{F}'(x) = (1/m) \sum_{j=0}^{m-1} (z_{[nx]+1-j} - \hat{\beta}_0 - \hat{\beta}_1([nx] + 1 - j)) \text{ if } nx \geq m - 1, \\ &\hat{F}'(x) = 0 \text{ if } [nx] < m - 1. \end{aligned} \tag{4b}$$

as the common cotrending vectors.

It is assumed that M_1 and M_2 have the same rank and

$$n^{-2p}\hat{M}_1 \to M_1 \tag{5a}$$

$$n^{-2p}\hat{M}_2 \to M_2 \tag{5b}$$

in probability for some non-negative number p. Therefore, the test statistic $n^{1-\alpha}\hat{\lambda}_r$ of $\Omega(r)$ against $\Omega(0)$ is based on the optimum solution $\hat{\lambda}_r$, say, that optimizes

TABLE I
Summary Statistics of Spot and Forward Rates

	Spot Rate, S_t			*Forward Rate, F_t*		
	C\$	*BP*	*SF*	*C\$*	*BP*	*SF*
Minimum	−0.468	0.052	−1.187	−0.468	0.048	−1.187
Maximum	0.038	0.949	−0.125	0.038	0.948	−0.122
Mean	−0.221	0.537	−0.561	−0.222	0.535	−0.558
	(0.016)	(0.010)	(0.015)	(0.016)	(0.011)	(0.015)
Variance	0.016	0.027	0.064	0.016	0.027	0.063
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
Skewness	0.232	0.452	−0.564	0.239	0.449	−0.576
Kurtosis	−0.622	−0.012	−0.588	−0.612	−0.002	−0.564

the general eigenvalue problem $\det(\hat{M}_1 - \lambda\hat{M}_2) = 0$, where α is the parameter of asymptotic power.[8]

Also, the test allows for testing the hypothesis that the spot rate and forward rate are nonlinearly cotrended with a nonlinear cotrending vector $(1, -1)^T$. Let H be a 2×1 vector, and let λ^* be the maximum solution of the generalized eigenvalue problem $\det(H^T\hat{M}_1H - \lambda H^T\hat{M}_2H) = 0$. If H spans a subspace of cotrending vectors then $n^{1-\alpha}\hat{\lambda}$ converges in distribution to the maximum eigenvalue $\bar{\lambda}^*$ of the matrix $\int \overline{W}(X)\overline{W}(X)^T dx$, where $\overline{W}(X)$ is a detrended Wiener process.

DATA, MODEL SELECTION, AND UNIT ROOTS

The end-of-the month observations of spot and forward exchange rates are obtained from DRI database. The sample includes the spot and forward rates of the British pound, the Canadian dollar, and the Swiss franc all against the US dollar. The empirical period spans from July 1973 to December 2007. To avoid measurement error, we follow the sampling procedure of Bekaert and Hodrick (1992), using the exact delivery dates of the forward contracts. Table I shows summary statistics of the spot and forward rates. Note that the data are transformed to natural logarithms.

We begin by finding the most suitable model for each of the series by employing several different types of unit root tests. This does not only help to characterize the time series features of the data but also allows us to address

[8]Bierens (2000) test use Newey–and West (1987) type variance estimator of the long-run variance of the error term, with truncation lag $m = [n^\alpha]$, it is shown that the value of $\alpha = 0.5$ of the distribution term is optimal for the convergence of $\hat{M}_2$ to M_2.

the question of whether the data are better modeled with a linear deterministic trend, or a nonlinear deterministic trend, and whether there is a unit root in the processes. The first type, in which the unit root exists under the null against the alternative of stationary and linear trend stationary, includes the traditional Augmented Dickey-Fuller (ADF) test and the Phillips and Perron (1988) (PP) test. The Akaike Information criterion (AIC) is used to determine the optimal number of lags for all the tests. We also employ the nonparametric methodology of Breitung (2002) to test for unit root (with drift) against the alternative hypothesis of stationarity (linear trend stationarity). There are two advantages of the Breitung (Bn) test over the ADF and PP tests: The Monte Carlo simulations show that it is robust to structural break, and to model misspecifications since the asymptotic property is independent from the stationary component of the series. Also, the Higher Order Autocorrelation (HOAC) tests suggested by Bierens (1993) are used to test the same hypotheses. The B(n) and (HOAC) tests are described in the Appendix A.

Then, we employ the methods of Kwiatkowski, Phillips, Schmidt, and Shin (1992) (KPSS) and Bierens and Guo (1993) (BG) to test the null hypothesis of (linear trend) stationarity against the unit root (without drift) hypothesis. The KPSS tests use a New-West (1987) type variance estimator of the long-run variance of the error term, with truncation lag $m = [c.n^r]$, where $c > 0$ and $0 < r < 0.5$. The default values of c and r are $c = 5$, $r = 0.25$ of the distribution term. For BG type tests, the null distribution is that of the absolute value of a standard Cauchy variate, except the one denoting BG(4), which employs the variance type as that of KPSS with the same truncation lag length.

Finally, we employ Bierens (1997) tests where the unit root with drift hypothesis is tested against the alternative of nonlinear trend stationarity. The $t(m)$, $A(m)$, $F(m)$, and $T(m)$ tests suggested by Bierens are described in the Appendix A. As shown in the Appendix A, the tests are based on an Augmented Dickey-Fuller auxiliary regression with nonlinear deterministic trends captured by transformed Chebishev polynomials that are orthogonal to time. It should be noted, however, that a rejection of the null does not necessarily imply that the process is nonlinear trend stationary. For example, for the $t(m)$ test statistic, a right-sided rejection of the null only provides information on the alternative. In contrast, the model-free $T(m)$ has the power to distinguish among three hypotheses; a left-sided rejection of the null suggests linear trend stationarity; and a right-sided rejection points in the direction of nonlinear trend stationarity.

Tables II and III report the results of the first and second types stationarity tests of the spot and forward rates, respectively. The first type: ADF, PP, (D)HOAC, and Breitung tests fail to reject the null hypothesis of a unit root (with drift) across all currencies, with sole exception of the British pound where the ADF tests strongly reject the hypothesis of a unit root in favor of

TABLE II

Nonstationarity Tests with a Constant or a Constant and a Linear Trend for the Spot Exchange Rate

	Test Statistics			*Critical Regions*			
Tests	*C$*	*BP*	*SF*	*5%*	*10%*	*H0*	*HA*
ADF(1)	−1.0925	−3.006**	−2.0478	<−2.89	<−2.58	UR	CS
ADF(2)	−2.2729	−3.374**	−2.0622	<−3.40	<−3.13	URD	TS
PP(1)	−1.72	−10.17	−6.39	<−14.51	<−11.65	UR	CS
PP(2)	−7.91	−13.96	−9.50	<−21.78	<−18.42	URD	TS
HOAC(1, 1)	−0.46	−6.86	−5.33	<−14.00	<−11.20	UR	CS
HOAC(2, 2)	−3.22	−8.29	−5.41	<−15.70	<−13.1	UR	CS
DHOAC(1, 1)	−2.10	−9.57	−5.64	<−20.60	<−17.10	URD	TS
DHOAC(2, 2)	−2.58	−9.60	−8.43	<−22.40	<−18.9	URD	TS
Breitung(1)	0.06928	0.03002	0.08325	>0.0102	>0.01449	UR	CS
Breitung(2)	0.00859	0.00401	0.00481	>0.0035	>0.00445	URD	TS
BG(1)	531.606**	44.541**	22.489**	>12.706	>6.314	CS	UR
BG (2)	344.880**	42.183**	27.860**	>12.706	>6.314	CS	UR
BG (3)	46.360**	8.209*	7.067*	>12.706	>6.314	CS	UR
BG (4)	37.954**	6.283	5.8289	>12.706	>6.314	CS	UR
BG (5)	57.558**	61.522**	66.133**	>12.706	>6.314	TS	URD
BG (6)	2.534	3.3947	4.7708	>12.706	>6.314	TS	URD
KPSS(1)	1.226**	0.739**	1.101**	>0.463	>0.347	CS	UR
KPSS(2)	0.159**	0.139*	0.106	>0.146	>0.119	TS	URD

Note. ADF(1), Augmented Dickey-Fuller test type (1); PP(1), Phillips-Perron test, type (1); (D) HOAC(1, 1,), Bierens' (detrended) higher order autocorrelation test type 1, 1; Breitung(1), Breitung test type 1; BG(1), Bierens-Guos' tests type 1; KPSS(1), Kwiatkowski, Phillips, Schmidt, and Shin's tests type 1. UR, unit root; CT, constant stationarity; TS, trend stationarity. * and ** denote rejection of the null at 10 and 5% significance level, respectively.

stationarity. Since these conventional tests lack power against the alternative hypothesis of stationarity, we use the more powerful KPSS and BG tests to verify these results. Unlike the conventional tests, some tests reject unit root and some reject stationarity; the BG(6) test does not reject linear trend stationarity for all the currencies, BG(3), BG(4), and KPSS(2) favors (linear trend) stationarity hypothesis for both the British pound and the Swiss franc, while all other tests favor unit root hypothesis. An elucidation of these contradictory results might be that the time series concerned are nonlinear trend stationary, with more complexities than the broken linear trend suggested by Perron (1989). For a closer examination of the nonstationarity of the spot and forward rates, we consider the more general trend stationarity alternative hypothesis developed by Bierens (1997).

The results of the nonlinear trend stationarity tests are shown in Table IV. The number of lagged first differences is set on the basis of Akaike criterion and the Chebishev time polynomial order $m = 10$. Since the tests are subject to size distortion, the table also reports the simulated p-values of the tests based

TABLE III
Nonstationarity Tests with a Constant or a Constant and a Linear Trend for the Forward Exchange Rate

	Test Statistics			*Critical Regions*			
Tests	*C$*	*BP*	*SF*	*5%*	*10%*	*H0*	*HA*
ADF(1)	−1.0971	−3.003*	−2.053	<−2.89	<−2.58	UR	CS
ADF(2)	−2.2672	−3.374*	−2.061	<−3.40	<−3.13	URD	TS
PP(1)	−1.76	−10.28	−6.49	<−14.51	<−11.65	UR	CS
PP(2)	−7.96	−14.06	−9.58	<−21.78	<−18.42	URD	TS
HOAC(1, 1)	−0.47	−6.85	−5.38	<−14.00	<−11.20	UR	CS
HOAC(2, 2)	−3.25	−8.28	−5.46	<−15.70	<−13.1	UR	CS
DHOAC(1, 1)	−2.13	−9.59	−5.67	<−20.60	<−17.10	URD	TS
DHOAC(2, 2)	−2.61	−9.62	−8.48	<−22.40	<−18.9	URD	TS
Breitung(1)	0.06928	0.08782	0.07001	>0.0102	>0.01449	UR	CS
Breitung(2)	0.00859	0.00723	0.00565	>0.0035	>0.00445	URD	TS
BG(1)	528.674**	42.736**	22.082**	>12.706	>6.314	CS	UR
BG (2)	344.612**	40.588**	27.093**	>12.706	>6.314	CS	UR
BG (3)	46.557**	8.090*	7.050*	>12.706	>6.314	CS	UR
BG (4)	38.134**	6.478*	5.776	>12.706	>6.314	CS	UR
BG (5)	61.949**	60.352**	64.443**	>12.706	>6.314	TS	URD
BG (6)	2.740	3.371	4.787	>12.706	>6.314	TS	URD
KPSS(1)	1.225**	0.737**	1.111**	>0.463	>0.347	CS	UR
KPSS(2)	0.158**	0.1389*	0.101	>0.146	>0.119	TS	URD

Note. ADF(1) = Augmented Dickey-Fuller test type (1); PP(1) = Phillips-Perron test, type (1); (D) HOAC(1, 1) = Bierens' (detrended) higher order autocorrelation test type 1, 1; Breitung(1) = Breitung test type 1, BG(1) = Bierens-Guos' tests type 1, KPSS(i) = Kwiatkowski, Phillips, Schmidt, and Shin's tests type 1. UR, unit root; URD, unit root with drift; CT, constant stationarity; TS, trend stationarity. * and ** denote rejection of the null at 10% and 5% significance level, respectively.

on the wild bootstrap.[9] The distributions of p-values are derived by performing the wild bootstrap with 10,000 replications under each hypothesis. As shown in the table, the exchange rates vary in their econometric behaviors. The p-values of $T(m)$ and $F(m)$ tests of the Canadian dollar and the British pound spot and forward rates point in the direction of nonlinear trend stationary. The $t(m)$ and $A(m)$ tests of the British pound have p-values below 5% which imply linear trend stationary.

However, one cannot reject the hypothesis that spot and forward rates of the Swiss franc are unit root processes. None of the tests are significant at 90% confidence level. This could contradict the results of BG and KPSS tests, which suggest linear trend stationarity. A reasonable justification is that the spot and forward rates of Swiss franc are neither genuine unit root nor genuine stationary.

[9]Refer to Mammen (1993) for a discussion of resampling methods. Härdle and Mammen (1993) suggest the wild bootstrap in a nonparametric regression.

TABLE IV
Bierens Tests of Unit Root with Drift Against Nonlinear Trend Stationarity

	Test Statistics			*Fractiles of the Asymptotic Null Distribution*			
Test Type	*C$*	*BP*	*SF*	*0.05*	*0.1*	*0.9*	*0.95*
Panel A: Nonstationarity tests of the spot exchange rates							
$t(m)$	−5.766	−6.415	−4.815	−6.67	−6.29	−3.86	−3.58
	−0.192	(0.034)*	−0.592				
$A(m)$	−69.614	−81.344	−51.5	−80.30	−73.7	−32.6	−29.6
	−0.162	(0.045)*	−0.446				
$F(m)$	4.543	4.8282	3.139	2.15	2.36	5.06	5.53
	(0.928)[a]	(0.956) **	−0.452				
$T(m)$	1778.047	1338.121	929.14	280.57	359.51	1660.07	1930.47
	(0.968)[b]	−0.836	−0.69				
Panel B: Nonstationarity tests of the forward exchange rates							
$t(m)$	−5.752	−6.405	−4.8057	−6.67	−6.29	−3.86	−3.58
	−0.22	(0.032)*	−0.676				
$A(m)$	−69.154	−80.999	−51.279	−80.30	−73.7	−32.6	−29.6
	−0.204	(.0420)*	−0.688				
$F(m)$	4.47	4.817	3.147	2.15	2.36	5.06	5.53
	(0.902)[a]	(0.962)[b]	−0.386				
$T(m)$	1715.579	1334.252	940.643	280.57	359.51	1660.07	1930.47
	(0.940)[a]	−0.822	−0.662				

Note. These tests are conducted two-sided. To adjust for small sample bias, the numbers in parenthesis are simulated *p*-values based on 1000 replications drown from the normal distribution with zero mean and OLS squared residuals variances (the wild bootstrap).

*Significant at 5% (rejection of unit root in favor to linear trend stationary).

**Significant at 5% (rejection of unit root in favor to linear trend stationary).

[a]Significant at 90% (rejection of unit root in favor to nonlinear trend stationary).

[b]Significant at 95% (rejection of unit root in favor to nonlinear trend stationary).

NONLINEAR COTRENDING BETWEEN FORWARD AND SPOT RATES

In this section, we present our empirical results. We begin by testing the null hypothesis that there are r cotrending vectors against the alternative that there are less than r cotrending vectors. We also estimate the relationship between the nonlinear trend component of the spot rate and the nonlinear trend component of the forward rate. Finally, we test the hypothesis that the spot and forward rates are nonlinearly cotrended with the cotrending vector $H = (1, -1)^T$.

Table V reports the test statistic $n^{1-\alpha}\hat{\lambda}_r$ and its corresponding 5 and 10% critical values for the vectors time series processes $z_{i,t} = (S_{i,t+1}, F_{it})^T$, where i stands for currency i. The results indicate that there is one cotrending vector between each currency forward rate and its corresponding future spot rate.

TABLE V
Tests of the Number of r Cotrending Vectors: $z_{i,t} = (S_{i,t+1}, F_{it})^T$

	Test Statistic			Critical Regions	
r	C$	BP	SF	5%	10%
1	0.14496	0.5713	0.06586	>0.46577	>0.35182
2	1.59868*	0.92262*	1.37568*	>0.67420	>0.53561

*Significant at 5%.

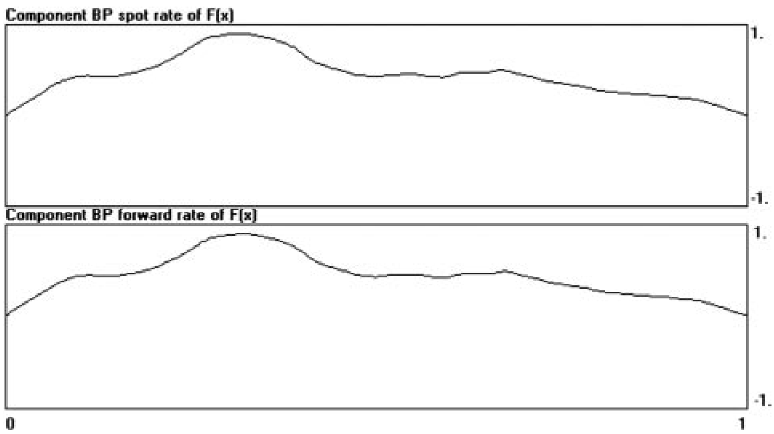

FIGURE 1
Rescaled component of $\hat{F}(X)$ of the British pound.

However, the hypothesis that there are two cotrending vectors is rejected across all currencies.

A unique cotrending vector between S_{t+1} and F_t entails a single-trend sharing. Figures 1–6 plot the components functions $\hat{F}(X)$ and $\hat{F}'(X)$, respectively, for the British pound, the Canadian dollar, and the Swiss franc. As shown, the common patterns in these functions obviously confirm the test results of the presence of nonlinear cotrending between each currency forward rate and its corresponding future spot rate.

Testing the hypothesis that the vector $H = (1, -1)^T$ is the cotrending vector is of particular interest, because it entails that the forward rate has a constant conditional expected value. Consider Fama (1984) decomposition of the forward rate

$$F_t = E_t[S_{t+1}] + P_t, \tag{6}$$

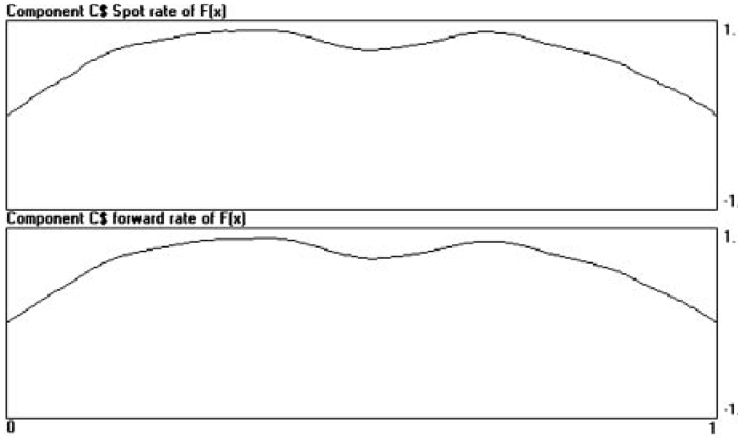

FIGURE 2
Rescaled component of $\hat{F}(X)$ of the Canadian dollar.

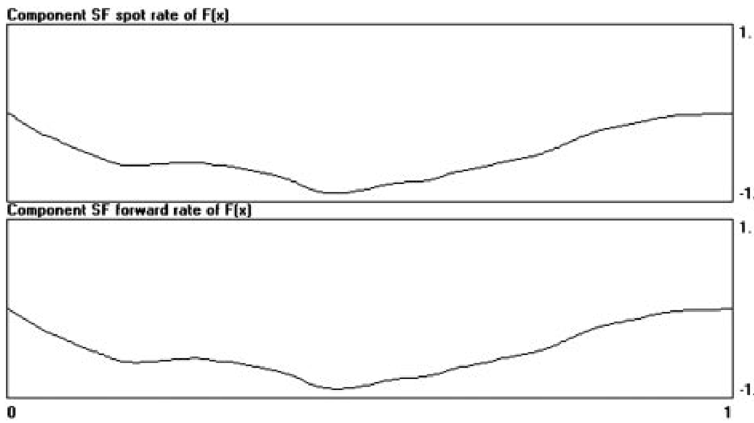

FIGURE 3
Rescaled component of $\hat{F}(X)$ of the Swiss franc.

where $E_t[\cdot]$ is the operator of rational expectations conditional on time's t information and P_t is the risk premium.

Under one form of rational expectations theory of forward rates unbiasedness, the premium P_t is a constant and time-invariant. Under even a stronger form of expectations theory, P_t does not exist for all t and the forward rate equals the expected value of the spot rate.

Table VI reports the cotrending vectors and the Bierens (2000) test of the hypothesis that the vector $H = (1, -1)^T$ is the cotrending vector of S_{t+1} and F_t.

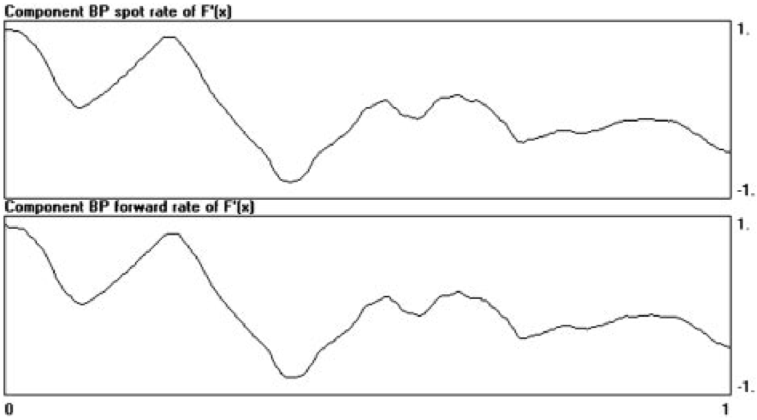

FIGURE 4
Rescaled component of $\hat{F}'(X)$ of the British pound.

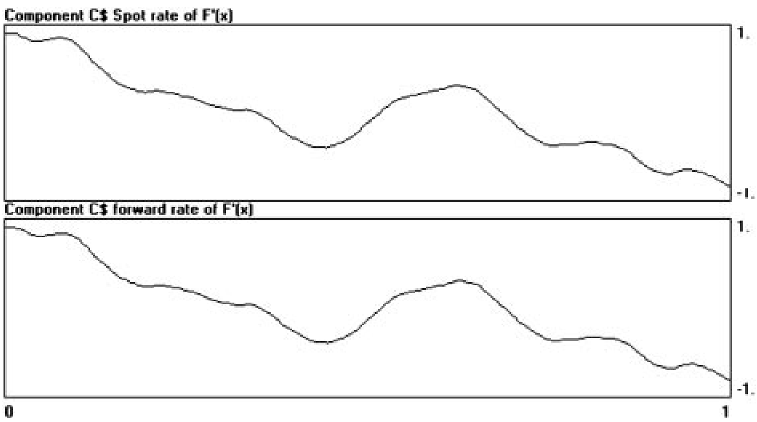

FIGURE 5
Rescaled component of $\hat{F}'(X)$ of the Canadian dollar.

As shown, the expectations theory is robust across all currencies. The $\overline{\lambda}^*$ max tests suggest that the forward rates of the British pound, the Canadian dollar, and the Swiss franc have constants expected values. All four currencies have $\overline{\lambda}^*$ max tests less than 0.352 and cannot be rejected at 95% confidence level. Thus we conclude that not only a robust nonlinear cotrending relation between S_{t+1} and F_t exists but also the unbiasedness of forward rate is sustained.

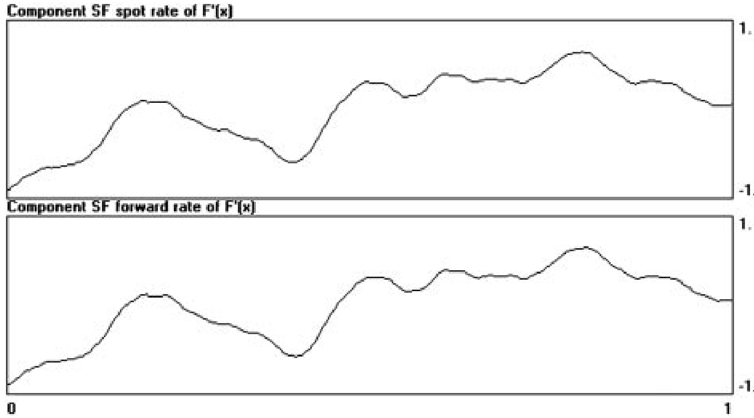

FIGURE 6
Rescaled component of $\hat{F}'(X)$ of the Swiss franc.

TABLE VI
Bierens Nonlinear Cotrending Vectors Between S_{t+1} and F_t and Tests of Time-Varying Risk Premia

Cotrending Vector	$\bar{\lambda}^*$ *max test*
Nonlinear trend in C$ S_{t+1} = 0.9974 × Nonlinear trend in C$ F_t	0.15
Nonlinear trend in BP S_{t+1} = 0.9919 × Nonlinear trend in BP F_t	0.07
Nonlinear trend in SF S_{t+1} = 0.9963 × Nonlinear trend in SF F_t	0.07

DOES FORWARD RATE UNBIASEDNESS HOLD

The results of Section 4 suggest that the nonlinear deterministic trend of F_t is efficient in tracking the nonlinear deterministic trend of S_{t+1}. If so, it is then possible that the stationary components (not the nonlinear trend components) of spot and forward rates are responsible for the routine rejection of the expectations hypothesis in the conventional test that reads:

$$\begin{aligned} S_{t+1} &= \alpha + \beta F_t + \varepsilon_{t+1} \\ H_0{:}\alpha &= 0, \beta = 1. \end{aligned} \tag{7}$$

In Table VII, we illustrate the conventional forward rate unbiasedness test on the basis of equation (7). Parameters estimate, p-values and Wald statistics for testing the joint hypothesis H_0: $\alpha = 0$, $\beta = 1$ are given in the table.[10]

[10]Consistent with Phillips, McFarland and McMahon (1996), the forward rate unbiasedness cannot be rejected across all currencies using the OLS estimator. Note that t-ratios and Wald statistics are not asymptotically valid because of the nonstationarity and temporal dependence of the data for the rationales explained in Phillips (1986), Park and Phillips (1988), and Deng (2005).

TABLE VII
The Conventional Tests of Unbiasedness of F_t in Predicting $S_{t+1} = \alpha + \beta F_t + \varepsilon_{t+1}$

	α	β	*Wald Statistic*
C$	−0.00078	0.9992	0.02
	(0.53575)	(0.88211)	(0.64323)
BP	0.01226	0.97767	5.13
	(0.02697)	(0.02390)	(0.07674)
SF	−0.00912	0.98611	3.99
	(0.05723)	(0.05968)	(0.13598)

Figures within parenthesis are p-values computed under the nulls that $\alpha = 0$, $\beta = 1$.

Wald statistics for the hypothesis H_0: $\alpha = 0$, $\beta = 1$.

To examine our hypothesis, we follow the detrending procedure of Hamming (1973) and Bierens (1997). Chebishev time polynomials of the form

$$P_{0,n}(t) = 1, P_{j,n}(t) = \sqrt{2}\cos[j\pi(t - 0.5)/n], j = 1, \ldots, n - 1. \tag{8}$$

which are orthogonal to time, are added to equation (7). The new model reads:

$$S_{t+1} = \alpha + \beta F_t + \sum_{j=0}^{n-1} \gamma_{j,n} P_{j,n}(t) + \varepsilon_{t+1} \tag{9}$$

where $\sum_{j=0}^{n-1} \gamma_{j,n} P_{j,n}(t)$ is the nonlinear trend function $g(t)$ and $\gamma_{j,n}$ is the parameter of decreasing smoothness $(1/n) \sum_{j=0}^{n-1} g(t) P_{j,n}(t)$. The results of such a specification are reported in Table VIII. As shown, the Wald statistics for testing the hypothesis that $\sum_{j=0}^{n-1} \gamma_{j,n} P_{j,n}(t) = 0$ are large and beyond 5% level, confirms Section 2's evidence that the spot and forward rates are nonlinear trend stationary processes. As expected, the insertion of Chebishev time polynomials to equation (7) sharply reduces the parameter estimates of β. Therefore, the hypothesis that the stationary components of the spot and forward rates account for the rejection of the forward rate unbiasedness in the conventional test is justified.[11]

To display evidence in support of this conclusion, we decompose the forward rate and its corresponding spot rate into there stationary and trend components. The resulting series are then used to compare the predictive power of each forward rate's component in forecasting its corresponding spot rate's component.

[11]The results are consistent with Phillips et al. (1996) and Phillips and McFarland (1997) that the OLS estimate with regards to the slope coefficient β is upward biased. Using the robust FM-LAD estimator, they report smaller values of β and they strongly reject the expectations hypothesis.

TABLE VIII

Tests of Unbiasedness of F_t in Predicting S_{t+1} with 10 Chebishev Polynomials Model: $S_{t+1} = \alpha + \beta F_t + \sum_{j=0}^{n-1} \gamma_{j,n} P_{j,n}(t) + \varepsilon_{t+1}$

	α	β	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}	*Wald Statistics*
C$	−0.0407	0.8188	0.0187	0.0001	0.0120	0.0015	−0.0030	0.00096	0.00467	−0.0015	−0.0027	−0.0004	64.48
	(0.0000)	(0.0000)	(0.0003)	(0.7990)	(0.0000)	(0.0110)	(0.0000)	(0.1757)	(0.0000)	(0.0473)	(0.0002)	(0.5433)	(0.00)
BP	0.0868	0.8384	0.0143	0.0071	0.0095	−0.0012	−0.0076	−0.0047	0.0100	0.0071	0.0074	0.0025	47.83
	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0010)	(0.0000)	(0.0020)	(0.0000)	(0.0000)	(0.0000)	(0.1040)	(0.00)
SF	−0.0662	0.8838	−0.0210	−0.0083	0.0069	−0.0073	−0.0069	−0.0091	−0.0029	0.0047	0.0016	0.0002	33.05
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0110)	(0.0030)	(0.0000)	(0.1360)	(0.0340)	(0.3920)	(0.8970)	(0.00)

Figures within parenthesis are *p*-values computed under the nulls that $\alpha = 0$, $\beta = 1$, and $\theta_i = 0$.

Wald statistics for the hypothesis that the Chebishev polynomials are jointly equal zero.

We begin by giving a brief clarification of the decomposition method. Suppose we write S_{t+1} and F_t as a sum of two components:

$$S_{t+1} = g_S(t + 1) + \xi_{t+1} \tag{10}$$

and

$$F_t = g_F(t) + \nu_t. \tag{11}$$

Where ξ_{t+1} and ν_t are the stationary components of the spot rate and forward rate, respectively. $g_S(t + 1)$ and $g_F(t)$ are trends functions of the spot rate and the forward rate, respectively. To arrive at these decompositions, each currency spot (forward) rate is regressed on 10 Chebishev time polynomials (with intercept). The OLS residuals of the regression are considered as a stationary component of the spot (forward) rate. These Chebishev polynomials $\left(\sum_{j=0}^{n-1} \gamma_{j,n} P_{j,n}(t) \right)$ are defined as a trend component of the spot (forward) rate.[12]

Test on unbiasedness of ν_t in predicting ξ_{t+1} can be based directly on the regression model

$$\xi_{t+1} = \kappa + \lambda \nu_t + \varepsilon_{t+1} \tag{12}$$

and the null hypothesis takes the parametric form

$$H_0: \kappa = 0, \lambda = 1.$$

While test on unbiasedness of $g_F(t)$ in predicting $g_S(t + 1)$ is concerned in testing the null hypothesis

$$\begin{gathered} g_S(t + 1) = \mu + \gamma g_F(t) + \varepsilon_{t+1} \\ H_0: \mu = 0, \gamma = 1. \end{gathered} \tag{13}$$

Table IX reports the estimation results of equation (12). As shown, the stationary components of the forward rates lack power in predicting the stationary components of the future spot rates. Looking at Table IX, we see that the value of parameter λ ranges between 0.81805 and 0.89233, which is sharply below the slope parameter β in the conventional test of the forward rates unbiasedness. The hypothesis that $\lambda = 1$ is strongly rejected across all currencies. The joint

[12]These Chebishev polynomials are jointly significant at 5% level for all spot (forward) exchange rates, confirming Section 3 evidence that the spot and forward rates are nonlinear trend stationary. The results are available on request.

TABLE IX

Tests of Unbiasedness of ν_t in Predicting ξ_{t+1} Model: $\xi_{t+1} = \kappa + \lambda v_t + \varepsilon_{t+1}$

	κ	λ	*Wald statistic*
Full sample			
C$	0.00003	0.81805	31.22
	(0.96777)	(0.0000)	(0.0000)
BP	−0.00029	0.89233	16.09
	(0.86399)	(0.0000)	(0.0003)
SF	−0.00002	0.88376	17.54
	(0.99253)	(0.0000)	(0.0001)
Sub-sample			
C$	−0.00046	0.81367	13.07
	(0.69093)	(0.0000)	(0.0000)
BP	−0.00047	0.85083	17.29
	(0.81242)	(0.0358)	(0.0003)
SF	0.00003	0.89358	11.14
	(0.77046)	(0.0027)	(0.0008)

Figures within parenthesis are *p*-values computed under the nulls that $\kappa = 0$, $\lambda = 1$.

Wald statistics for the hypothesis H0: $\kappa = 0$, $\lambda = 1$.

TABLE X

Tests of Unbiasedness of $g_F(t)$ in Predicting $g_S(t+1)$ Model: $g_s(t+1) = \mu + \gamma\, g_F(t) + \varepsilon_{t+1}$

	μ	γ	*Wald Statistic*
Full sample			
C$	0.00022	1.00418	1.52
	(0.96777)	(0.42117)	(0.46655)
BP	−0.02367	0.99405	0.35
	(0.99342)	(0.55597)	(0.82622)
SF	−0.00362	0.99607	0.74
	(0.99253)	(0.62406)	(0.62406)
Sub-sample			
C$	0.00596	1.02114	5.32
	(0.00072)	(0.31635)	(0.02111)
BP	0.00179	1.00123	0.14
	(0.27912)	(0.62917)	(0.71018)
SF	−0.00598	0.98623	31.77
	(0.00000)	(0.28216)	(0.0000)

Figures within parenthesis are *p*-values computed under the nulls that $\mu = 0$, $\gamma = 1$.

Wald statistics for the hypothesis H_0: $\mu = 0$, $\gamma = 1$.

hypothesis H_0: $\kappa = 0$, $\lambda = 1$ is also rejected. The same results can be gleaned using out-of-sample data (the last 120 observations of the data).

Table X reports the estimation results of equation (13). Consistent with the results in section 4, the nonlinear trend component of the forward rate is an unbiased predictor of the nonlinear trend component of the future spot rate.

The Wald statistics for the hypothesis H_0: $\mu = 0$, $\gamma = 1$ cannot be rejected at any plausible level of significance. All currencies have χ^2 P-values in excess of 0.4 and the hypothesis cannot be rejected at even the 90% confidence level. For the out-of-sample data, the results are mixed. The hypothesis that $\gamma = 1$ cannot be rejected across all currencies but the joint hypothesis is rejected for both Canadian dollar and Swiss franc.

WHY FORWARD PREMIA MOVE IN THE OPPOSITE DIRECTION WITH SPOT CHANGES

The results of Sections 4 and 5 suggest that the positive co-movement between the forward rate and its corresponding spot rate is due mainly to a common nonlinear deterministic trend. In contrast, the stationary component of the forward rate understate the stationary component of the future spot rate. There is a good reason therefore to question whether the downward bias prediction of the stationary component of the forward rate is responsible for the forward premium puzzle? Note that the forward premium puzzle is that the subsequent spot changes $S_{t+1} - S_t$ are negatively related to the forward premium $F_t - S_t$, whereas the expectations theory foretells that the spot rate will decline when the forward sells at discount.

Following Fama (1984), the forward premium puzzle test can be based directly on the regression model

$$S_{t+1} - S_t = \alpha + \beta(F_t - S_t) + \varepsilon_{t+1} \tag{14}$$

and the expectations hypothesis takes the parametric form

$$H_0: \alpha = 0, \beta = 1.$$

Table XI illustrates the test on the basis of equation (14). Parameter estimates, p-values, and Wald statistics for testing the joint hypothesis H_0: $\alpha = 0$, $\beta = 1$ are given in the table.[13]

To begin the exploration of whether the downward bias of the stationary component of the forward rate (in predicting the stationary component of the spot rate) is responsible for the forward premium puzzle, suppose the true process generating $g_S(t + 1)$ is

$$g_S(t + 1) = \theta g_S(t) + \varsigma_{t+1} \tag{15}$$

[13]Consistent with Fama (1984), Bilson (1981), and Gregory and McCurdy (1984), it can be seen that the rejection of the hypothesis is overwhelming.

TABLE XI

The Conventional Tests of the Forward Premium Model: $S_{t+1} - S_t = \alpha + \beta(F_t - S_t) + \varepsilon_{t+1}$

	α	β	*Wald Statistic*
C$	−0.00014	−1.31242	15.03
	(0.82877)	(0.88211)	(0.00055)
BP	0.00016	−1.71111	14.44
	(0.91972)	(0.0239)	(0.00073)
SF	0.00034	−4.26602	18.10
	(0.85592)	(0.05968)	(0.00023)

Figures within parenthesis are p-values computed under the nulls that $\alpha = 0$, $\beta = 1$.

Wald statistics for the hypothesis H_0: $\alpha = 0$, $\beta = 1$.

where $E_t(\varsigma_{t+1}) = 0$. The true process generating ξ_{t+1} is

$$\xi_{t+1} = \vartheta \xi_t + \varsigma'_{t+1} \tag{16}$$

where $E_t(\varsigma'_{t+1}) = 0$ and $\vartheta < 1$ because ξ_{t+1} is stationary.

Consider the least squares estimator of β in equation (14):

$$\hat{\beta} = \frac{\sum_{t=1}^{n} (F_t - S_t)(S_{t+1} - S_t)}{\sum_{t=1}^{n} (F_t - S_t)^2} \tag{17}$$

where n denotes sample size. Decomposing F_t and S_t to their stationary and nonlinear trend components and substituting into the formula for $\hat{\beta}$ yields,

$$\hat{\beta} = \frac{\sum_{t=1}^{n} (g_F(t) + \nu_t - g_s(t) - \xi_t)(g_s(t+1) + \xi_{t+1} - g_s(t) - \xi_t)}{\sum_{t=1}^{n} (g_F(t) + \nu_t - g_s(t) - \xi_t)^2} \tag{18}$$

Consider the empirical findings of sections 4 and 5 where the lines of best fit read:

$$g_S(t+1) = g_F(t), \tag{19}$$

and

$$\xi_{t+1} = \lambda \nu_t, \tag{20}$$

where $\lambda < 1$. Typically, we find that $g_F(t)$ is an unbiased predictor of $g_S(t+1)$ and ν_t is a downward biased predictor of ξ_{t+1}.

Substituting this and equations (15), (16), into (18) yields,

$$\hat{\beta} = \frac{\sum_{t=1}^{n}(g_s(t+1) + \xi_{t+1}/\lambda - g_s(t) - \xi_t)(g_s(t+1) + \xi_{t+1} - g_s(t) - \xi_t)}{\sum_{t=1}^{n}(g_s(t+1) + \xi_{t+1}/\lambda - g_s(t) - \xi_t)^2} \tag{21}$$

$$\hat{\beta} = \frac{\sum_{t=1}^{n}\left(\theta g_S(t) + \frac{\vartheta}{\lambda}\xi_t - g_s(t) - \xi_t\right)(\theta g_S(t) + \text{s}_{t+1} + \vartheta\xi_t + \text{s}'_{t+1} - g_s(t) - \xi_t)}{\sum_{t=1}^{n}\left(\theta g_S(t) + \frac{\vartheta}{\lambda}\xi_t - g_s(t) - \xi_t\right)^2} \tag{22}$$

which can be written as

$$\hat{\beta} = 1 - \frac{\sum_{t=1}^{n}\left[(\theta-1)g_S(t) + \left(\frac{\vartheta}{\lambda} - 1\right)\xi_t\right]\left[\left(\frac{\vartheta}{C} - \vartheta\right)\xi_t - \text{s}_{t+1} - \text{s}'_{t+1}\right]}{\sum_{t=1}^{n}\left((\theta-1)g_S(t) + \left(\frac{\vartheta}{\lambda} - 1\right)\xi_t\right)^2} \tag{23}$$

Under forward-rate unbiasedness hypothesis, $\beta = 1$, and so

$$\hat{\beta} - \beta = -\frac{\sum_{t=1}^{n}\left[(\theta-1)g_S(t) + \left(\frac{\vartheta}{\lambda} - 1\right)\xi_t\right]\left[\left(\frac{\vartheta}{\lambda} - \vartheta\right)\xi_t - \text{s}_{t+1} - \text{s}'_{t+1}\right]}{\sum_{t=1}^{n}\left((\theta-1)g_S(t) + \left(\frac{\vartheta}{\lambda} - 1\right)\xi_t\right)^2}. \tag{24}$$

Expanding the expression and taking the probability limits yields

$$p\lim(\hat{\beta}) - \beta =$$

$$-\frac{\left[\begin{array}{l}(\theta-1)\left(\left(\frac{\vartheta}{\lambda} - \vartheta\right)\text{cov}(g_S(t), \xi_t) - \text{cov}\,(g_S(t), \text{s}_{t+1}) - \text{cov}(g_S(t), \text{s}'_{t+1}) + \right. \\ \left.\left(\frac{\vartheta}{\lambda} - 1\right)\left(\frac{\vartheta}{\lambda} - \vartheta\right)\text{var}(\xi_t) - \text{cov}(\xi_t, \text{s}_{t+1}) - \text{cov}(\xi_t, \text{s}'_{t+1})\right)\end{array}\right]}{\left[\begin{array}{l}(\theta-1)^2(\text{var}(g_S(t)) + (\overline{g}_S(t))^2) + \\ 2\left((\theta-1) + \left(\frac{\vartheta}{\lambda} - 1\right)\right)\text{cov}(g_s(t), \xi_t) + \left(\frac{\vartheta}{\lambda} - 1\right)^2\text{var}(\xi_t)\end{array}\right]}$$

If the stationary and nonlinear trend components of exchange rates are uncorrelated in the sample

$$p\lim(\hat{\beta}) - \beta = -\frac{\left(\frac{\vartheta}{\lambda} - 1\right)\left(\frac{\vartheta}{\lambda} - \vartheta\right)\mathrm{var}(\xi_t)}{(\theta - 1)^2(\mathrm{var}(g_S(t)) + (\bar{g}_S(t))^2) + \left(\frac{\vartheta}{\lambda} - 1\right)^2\mathrm{var}(\xi_t)} \tag{25}$$

Expression (27) highlights the result that the probability limit of $\hat{\beta}$ differs from unity in proportion to the variance of the stationary component of the spot rate to the sum of the squared mean, $(\bar{g}_S(t))^2$, and the variance, $var(g_S(t))$, of the nonlinear trend component of the spot rate. According to the expression, the deviation of $p\lim(\hat{\beta})$ from unity is governed primarily by the autocorrelation parameter ϑ and the slope of the regression of the stationary component of the forward rate on the stationary component of the future spot rate λ. If $\vartheta = \lambda$, the $p\lim(\hat{\beta}) = \beta$ because $\left(\frac{\vartheta}{\lambda} - 1\right) = 0$. If the stationary component of the forward rate fully predicts the stationary component of the future spot rate, that is $\lambda = 1$, then the "bias from unity" is also zero because $\left(\frac{\vartheta}{\lambda} - \vartheta\right) = 0$. Finally, the negative values observed for $\hat{\beta}$ must imply that $\vartheta > \lambda$. This implies that there is a stronger correlation between the stationary components of the future and current spot rate than between the stationary components of the current spot and forward rates.

To verify these predictions about the deviation of $\hat{\beta}$, we run Monte Carlo sampling experiments in which the stationary and trend components of the spot rate were assumed to be generated by the parameter estimates of equations (15) and (16) (with intercepts).

To generate pseudo-data on ξ_{t+1} and $g_S(t + 1)$ we need first to generate innovations ς_{t+1} and ς'_{t+1} in autoregressions (16) and (17). To do so, we employ the bootstrap approach suggested by Efron (1979) which produces innovations ς_{t+1} and ς'_{t+1} drawn randomly from the sample distributions of the errors ς'_{t+1} and ς'^{*}_{t+1} in the estimated autoregressions (16) and (17). The bootstrap samples $\{g^*_s(t+1): t = 1, \ldots n\}$ and $\{\xi^*_{s+1}:\ t = 1, \ldots n\}$ are generated from $g^*_s(t+1) = \theta g_S(t) + \varsigma^*_{t+1}$ and $\xi^*_{t+1} = \vartheta\xi_t + \varsigma'^{*}_{t+1}$, respectively, where $\{\varsigma^*_{t+1}: t = 1, \ldots, n\}$ and $\{\varsigma'^{*}_{t+1}: t = 1, \ldots, n\}$ are random draws with replacement from the empirical distributions of centered versions of innovations $\{\varsigma_{t+1} = g_S(t + 1) - \theta g_S(t): t = 1, \ldots, n\}$ and $\{\varsigma'_{t+1} = \xi_{t+1} - \vartheta\xi_t: t = 1, \ldots, n\}$, respectively. For the forward rates, two simulations were performed. The first, where the stationary component of the forward rates were generated recursively through the autoregression, $\nu_t = \omega + \phi\nu_{t-1} + \iota_t$. The second, where the stationary component of the forward rate are set to the rational forecast, $\nu_t = \omega' + \phi'\xi_t + \iota'_t$, of ξ_{t+1}. In both the simulations, the nonlinear trend components of the forward rates were generated recursively through the autoregression

TABLE XII
Sampling Distributions of $\hat{\beta}$

	$\hat{\beta}$-RECURSIVE			$\hat{\beta}$-RATIONAL		
Percentile (%)	*C$*	*BP*	*SF*	*C$*	*BP*	*SF*
100 (maximum)	0.797	1.481	1.539	1.00307	1.04563	0.67482
99	0.162	0.018	0.340	0.96673	1.00598	0.60015
95	−0.113	−0.518	−0.133	0.95271	0.99048	0.56942
90	−0.264	−0.780	−0.389	0.94457	0.98150	0.55166
75	−0.501	−1.188	−0.833	0.93153	0.96681	0.52451
50 (median)	−0.770	−1.641	−1.340	0.91670	0.95011	0.49425
25	−1.048	−2.080	−1.833	0.90173	0.93397	0.46405
10	−1.290	−2.462	−2.266	0.88789	0.91962	0.43697
5	−1.426	−2.672	−2.537	0.87947	0.91125	0.42194
1	−1.683	−3.028	−3.033	0.86447	0.89391	0.39229
0 (minimum)	−2.237	−4.103	−3.799	0.82005	0.86223	0.32135

10,000 repetitions were used to generate each distribution.

$g_f(t) = \omega' + \phi' g_f(t-1) + \iota'_t$. The simulations were performed using a sample size of $T = 414$ (the original sample size) and 10,000 repetitions.

The methodology presently explained is used to generate $T + 1$ pseudo-observations on ξ_t, ν_t, $g_S(t)$, and $g_F(t)$ and the coefficient estimates of $\hat{\alpha}$ and $\hat{\beta}$ were obtained for the regression in equation 14 using the pseudo-data.

Table XII contains summary statistics on the sampling distribution of $\hat{\beta}$ for each of the two simulations. The first set of distributions, labeled $\hat{\beta}$-RECURSIVE, reports $\hat{\beta}$ distributions with recursive generation of the stationary components of the forward rates. The first set of results comprises the evidence of forward premium puzzle referred to in the literature. The second set of distributions, labeled $\hat{\beta}$-RATIONAL, reports $\hat{\beta}$ distributions where the stationary components of the forward rates are set to the rational forecast. Both of the simulations have the general characteristics expected by the analysis above.

The most critical finding in Table XII is that after setting the stationary component of the forward rates to the rational forecast, the evidence against positive correlation between the premium and expected future spot rate components of forward rates disappears. This finding can be best illustrated by looking at the sampling distribution of $\hat{\beta}$-RATIONAL in panel B. The median estimates of $\hat{\beta}$ obtained in the 10,000 replications are 0.91670 for the Canadian dollar, 0.95011 for the British pound, and 0.49425 for the Swiss franc. The observed $\hat{\beta}$ of the currencies are 75% of the time above 0.9315, 0.96681, and 0.52451, with 99% of the time above 0.86447, 0.89391, and

0.39229, respectively. Therefore, the conclusion seems justified that the forward premium puzzle is due mainly to the biasedness of the transitory component of the forward rate in predicting the transitory component of the future spot rate.

CONCLUSIONS

Our goal in this paper is to understand the interaction between the forward rate and its corresponding future spot rate. To do this, we decompose the spot and forward rates into (permanent) nonlinear trend components and (transitory) stationary components. We then proceed to determine whether spot rate moves predominantly with the transitory or the permanent components of the forward rate.

Using new functional techniques, we provide statistical evidence of nonlinear deterministic trend behavior for the spot and forward exchange rates. We argue that the rejection of the forward rate unbiasedness hypothesis to the failure of the transitory component of the forward rate to fully predicts the transitory component of the future spot rate. We conclude that the forward rate is poor in tracking spot rate movements over short horizons. However, the permanent component of the forward rate is an unbiased predictor of the permanent component of the future spot rate. A robust nonlinear cotrending relation is also found between the forward and future spot rates and the hypothesis of the forward rate unbiasedness is sustained in the long run. These results would seem to help explain the long-standing difficulty of detecting a short-run relationship between exchange rates and macroeconomic measurable fundamentals (e.g. Mark, 1995; Meese & Rogoff, 1983).

Within these broad-based results, we also find an interesting explanation for the forward premium puzzle. We proceed by estimating Fama (1984) regression by simulating two sets of forward and spot rates. The first set, where the transitory components of the forward rates, were generated recursively through the autoregression implied from the estimated transitory component. In the second set, the transitory components of the forward rates were set to the rational forecast of the stationary component of the future spot rate. In both simulations, the spot rates components and the permanent components of the forward rates are generated recursively. We find that after setting the stationary component of the forward rates to the rational forecast, the evidence against positive correlation between the premium and expected future spot rate components of forward rates disappears. This suggests that the transitory components of the spot and forward rates are responsible for the forward premium puzzle.

APPENDIX

The Breitung (2002) nonparametric unit root test is based on the following variance ratio: let $R(t)$ be the partial sums of the spot rate

$$R(t) = s(1) + s(2) + \ .....\ + s(t)$$

The variance ratio, denoted by $B(n)$, can be computed based on the following formula:

$$B(n) = \frac{[R(1)^2 + R(2)^2 + \ .....\ + R(n)^2]n^{-1}}{[s(1)^2 + s(2)^2 + \ .....\ + s(n)^2]}$$

Under the unit root hypothesis $B(n)/n$ converges in distribution to a function of a standard Wiener process, which is free of nuisance parameters. On the other hand, if $s(t)$ is non-zero mean stationary then the demeaned version of $B(n)$ itself converges in distribution, hence $B(n)/n$ converges in probability to zero. If the alternative hypothesis is that $r(t)$ is linear trend stationary, then the detrended version of $B(n)/n$ converges to zero.

BIBLIOGRAPHY

Abuaf, N., & Jorion, P. (1990). Purchasing power parity in the long run. Journal of Finance, 45, 157–174.

Alder, M., & Lehmann, B. (1983). Deviation from purchasing power parity in the long run. Journal of Finance, 38, 1471–1487.

Ang, A., Bekaert, G., & Wei, M. (2008). The term structure of real rates and expected inflation. Journal of Finance, 63, 797–849.

Baldwin, R. (1990). Re-interpreting the failure of foreign exchange market efficiency tests: Small transaction costs, big hysteresis bands (working paper). National Bureau of Economic Research.

Balke, N. S., & Fomby, T. B. (1997). Threshold cointegration. International Economic Review, 38, 627–645.

Bams, D., Walkowiak, K., & Wolff, C. C. P. (2004). More evidence on the dollar risk premium in the foreign exchange market. Journal of International Money and Finance, 23, 271–282.

Bansal, R. (1997). An exploration of the forward premium puzzle in currency markets. Review of Financial Studies, 10, 369–403.

Bansal, R., & Dahlquist, M. (2000). The forward premium puzzle: Different tales from developed and emerging economics. Journal of International Economics, 51, 115–144.

Bekaert, G., & Hodrick, R. (1992). Characterizing predictable components in excess returns on equity and foreign exchange markets. Journal of Finance, 47, 467–510.

Berk, J., Green, R. C., & Naik, V. (1999). Optimal investment, growth options, and security returns. Journal of Finance, 54, 1553–1607.

Bierens, H. J. (1993). Higher-order sample autocorrelations and the unit root hypothesis. Journal of Econometrics, 57, 137–160.

Bierens, H. J. (1997). Testing the unit root with drift hypothesis against nonlinear trend stationary, with an application to the US price level and interest rate. Journal of Econometrics, 81, 29–64.

Bierens, H. J. (2000). Nonparametric nonlinear cotrending analysis, with an application to interest and inflation in the United States. Journal of Business and Economics Statistics, 18, 323–337.

Bierens, H. J., & Guo, S. (1993). Testing stationarity and trend stationarity against the unit root hypothesis. Econometric Reviews, 12, 1–32.

Bilson, J. F. (1981). The "speculative efficiency" hypothesis. Journal of Business, 54, 435–451.

Breitung, J. (2002). Nonparametric tests for unit roots and cointegration. Journal of Econometrics, 108, 343–364.

Campbell, J., & Cochrane, J. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of Political Economy, 107, 205–251.

Cavaglia, S., Verschoor, W., & Wolf, C. (1993). Further evidence of exchange rate expectations. Journal of International Money and Finance, 12, 78–98.

Chordia, T., & Shivakumar, L. (2002). Momentum, business cycle, and time-varying expected returns. Journal of Finance, 57, 985–1019.

Clarida, R., Gali, J., & Gertler, M. (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. Quarterly Journal of Economics, 115, 147–180.

Conrad, J., & Kaul, G. (1998). An anatomy of trading strategies. Review of Financial Studies, 11, 489–519.

Corbae, D., Lim, K., & Oulairis, K. (1992). On cointegration and tests of forward market unbiasedness. Review of Economics and Statistics, 74, 728–732.

Crowder, W. J. (1994). Foreign exchange market efficiency and common stochastic trends. Journal of International Money and Finance, 13, 551–564.

Deng, A. (2005). Understanding spurious regression in financial economics (working paper). Boston University.

Dumas, B. (1992). Dynamic equilibrium and the real exchange rate in a spatially separated world. Review of Financial Studies, 5, 153–180.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. Annals of Statistics, 7, 1–26.

Engle, C. (1996). The forward discount anomaly and the risk premium: A survey of recent evidence. Journal of Empirical Finance, 3, 123–192.

Engle, C., & Hamilton, J. (1990). Long swings in the dollar: Are they in the data and do markets know it? American Economic Review, 80, 689–713.

Engle, C., & West, K. D. (2004). Accounting for exchange rate variability in present value models when the discount factor is near one. American Economic Review, 94, 118–125.

Engle, C., & West, K. D. (2005). Exchange rates and fundamentals. Journal of Political Economy, 113, 485–517.

Epstein, L., & Zin, S. (1989). Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. Econometrica, 57, 937–969.

Evans, M. D., & Lyons, R. (2005). Meese-Rogoff redux: Micro-based exchange-rate forecasting. American Economic Review, 95, 405–414.

Fama, E. F. (1984). Forward and spot exchange rates. Journal of Monetary Economics, 14, 319–338.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33, 3–56.

Frankel, J. A. (1981). Flexible exchange rates, prices and the role of 'news': Lessons from the 1970's. Journal of Political Economy, 89, 665–705.

Frankel, J. A., & Froot, K. A. (1990). Chartists, fundamentalists, and trading in the foreign exchange market. American Economic Review, 80, 181–185.

Gregory, A. W., & McCurdy, T. H. (1984). Testing the unbiasedness hypothesis in the forward foreign exchange market: A specification analysis. Journal of International Money and Finance, 3, 357–368.

Hamming, R. W. (1973). Numerical methods for scientists and engineers. New York: Dover Publications.

Hansen, L. P., & Hodrick, R. J. (1983). Risk averse speculation in the forward exchange markets: An econometric analysis of linear models. In J. A. Frenkel (Ed.), Exchange Rates and International Economics. Chicago: University of Chicago Press.

Härdle, W., & Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. Annals of Statistics, 21, 1926–1947.

Hodrick, R. J. (1987). The empirical evidence on the efficiency of forward and futures foreign exchange markets. Chur: Harwood Academic Publishers.

Hollifield, B., & Uppal, R. (1997). An examination of uncovered interest rate parity in segmented international commodity markets. Journal of Finance, 52, 2145–2170.

Huizinga, J. (1987). An empirical investigation of the long-run behavior of real exchange rates; comment. Carnegie–Rochester Conference Series on Public Policy, 87, 225–231.

Hsieh, D. A. (1989). Testing for nonlinear dependence in daily foreign exchange rates. Journal of Business, 62, 339–368.

Kilian, L., & Taylor, M. (2003). Why is it difficult to beat the random walk forecast of exchange rates? Journal of International Economics, 60, 85–107.

Krasker, W. (1980). The peso problem in testing the efficiency of forward exchange markets. Journal of Monetary Economics, 6, 269–276.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null of stationarity against the alternative of a unit root. Journal of Econometrics, 54, 159–178.

Lettau, M., & Ludvigson, S. (2009). Measuring and modeling variation in the risk return tradeoff. In Y. Ait-Sahalia, & L. Peter Hansen (Eds.), Handbook of Financial Econometrics. Amsterdam: Elsevier.

Lewis, K. K. (1995). Puzzles in international financial markets. In G. Grossman, G., & K. Rogoff (Eds.), The Handbook of International Economics, Vol. III. Amsterdam: New York and Oxford: Elsevier, North-Holland.

Lyons, R. K. (2001). The Microstructure Approach to Exchange Rates. MA: MIT Press.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. Annals of Statistics, 21, 255–285.

Mark, N. C. (1995). Exchange rates and fundamentals: Evidence on long-horizon predictability. American Economic Review, 85, 201–218.

Mark, N. C., & Moh, Y.-K. (2004). Official interventions and occasional violations of uncovered interest parity in the dollar-dm market (working paper). Notre Dame University.

Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out-of-sample? Journal of International Economics, 14, 3–24.

Newey, W., & West, K. (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55, 703–708.

Nijman, T. E., Palm, F. C., & Wolff, C. C. P. (1993). Premia in forward foreign exchange as unobserved components. Journal of Business and Economic Statistics, 11, 361–365.

Nikoaou, K. (2008). The behavior of the real exchange rate: Evidence from regression quantiles. Journal of Banking and Finance, 32, 664–679.

Obstfeld, M., & Taylor, A. (1997). Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited. Journal of the Japanese and International Economies, 11, 441–479.

Park, J. Y., & Phillips, P. C. B. (1988). Statistical inference in regressions with integrated processes: Part 1. Econometric Theory, 4, 468–497.

Perron, P. (1989). The great crash, the oil-price shock, and the unit-root hypothesis. Econometrica, 57, 1361–1402.

Phillips, P. C. B. (1986). Understanding spurious regressions in econometrics. Journal of Econometrics, 33, 311–340.

Phillips, P. C. B., & McFarland, J. W. (1997). Forward exchange market unbiasedness: The case of the US dollar since 1984. Journal of International Money and Finance, 16, 885–907.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. Biometrica, 75, 335–346.

Phillips, P. C. B., McFarland, J. W., & McMahon, P. C. (1996). Robust tests of forward exchange market efficiency with empirical evidence from the 1920s. Journal of Applied Econometrics, 11, 1–23.

Rogoff, K. (1996). The purchasing power parity puzzle. Journal of Economic Literature, 34, 647–668.

Roll, R. (1979). Violations of purchasing power parity and their implications for efficient international commodity markets. In M. Sanat, & G. P. Szego (Eds.), International Finance and Trade, Vol. I, Cambridge, MA: Ballinger.

Sarno, L., Valente, G., & Leon, H. (2006). Nonlinearity in deviations from uncovered interest parity: An explanation of the forward bias puzzle. Review of Finance, 10, 443–482.

Sims, C. (1988). Bayesian skepticism on unit root econometrics. Journal of Economic Dynamics and Control, 12, 463–474.

Smallwood, A. D. (2008). Measuring the persistence of deviations from purchasing power parity with a fractionally integrated STAR model. Journal of International Money and Finance, 27, 1161–1176.

Sollis, R. (2008). U.S. dollar real exchange rates: nonlinearity revisited. Journal of International Money and Finance, 27, 516–528.
Summers, L. (1986). Does the stock market rationally reflect fundamentals values? Journal of Finance, 41, 591–601.
Taylor, M. P., & Peel, D. A. (2000). Nonlinear adjustment, long run equilibrium and exchange rate fundamentals. Journal of International Money and Finance, 19, 33–53.
Taylor, M. P., Peel, D. A., & Sarno, L. (2001). Nonlinear adjustment in real exchange rates: Towards a solution to the purchasing power parity puzzles. International Economic Review, 42, 1015–1042.
TerÄasvirta, T. (1994). Specification, estimation and evaluation of smooth transition autoregressive models. Journal of the American Statistical Association, 89, 208–218.
Verdelhan, A. (2010). A Habit-based explanation of the exchange rate risk premium. Journal of Finance, 65, 1, 123–145.
Whitt, J. (1992). The long-run behavior of the real exchange rate: A reconsideration. Journal of Money, Credit and Banking, 24, 72–82.
Wu, J., & Chen, S. (1998). Foreign exchange market efficiency revisited. Journal of International Money and Finance, 17, 831–838.
Yang, J., Su, X., & Kolari, J. W. (2008). Do Euro exchange rates follow a martingale? Some out-of-sample evidence. Journal of Banking and Finance, 32, 729–749.